



TITLE:

多項式の前処理によるPCG法とその応用(数値解析の基礎理論とその周辺)

AUTHOR(S):

野寺, 隆

CITATION:

野寺, 隆. 多項式の前処理によるPCG法とその応用(数値解析の基礎理論とその周辺). 数理解析研究所講究録 1988, 676: 137-147

ISSUE DATE:

1988-12

URL:

<http://hdl.handle.net/2433/100965>

RIGHT:

多項式の前処理による PCG 法とその応用

慶應義塾大学理工学部 野寺 隆 (Takashi Nodera)

概要. 対称な正定値行列を係数行列とする連立一次方程式の解法として不完全コレスキー分解を前処理とする共役勾配法は, 理工学の様々な分野で利用されるようになってきた. しかし, 対称ではあるが, 正定値でない行列に対して不完全コレスキー分解を利用することは, あまり得策であるとはいえない. そこで, 不完全コレスキー分解の代りに多項式近似による不完全逆行列 (incomplete inverse matrix) を前処理に利用する共役勾配法について考えることにする.

1. はじめに

偏微分方程式の差分近似から得られる大規模な線形方程式:

$$Ax = b \quad (1.1)$$

の近似解の計算に, 前処理付きの共役勾配法 (Preconditioned Conjugate Gradient method, または PCG 法とも言う) の利用は, 近年, 反復解法の中でも重要な位置をしめつつある. 特に, 前処理に不完全コレスキー分解を利用した ICCG 法 (Incomplete Cholesky Conjugate Gradient method) は, 加速パラメータを使用せずに共役勾配法を適切に加速し有限回の反復回数で収束するアルゴリズムとして広く一般に利用されている (Meijerink et al.[7]).

係数行列 A が対称で正定値の場合には, 共役勾配法は非常に有効なアルゴリズムである. しかし, 行列 A が対称ではあるが正定値でないときに

は、共役残差法 (CR 法) や通常、大型疎行列の固有値問題に適用されるランチョス法 (Lanczos method) に基づいた SYMLQ アルゴリズムも有力な解法である (Chandra[2]) .

もう一つのアプローチは昔から行われてきた方法であるが、(1) の方程式に係数行列 A を掛けて正規方程式 $A^2x = Ab$ を作り、これに共役勾配法を適用することである。しかし、残念ながら行列 A の条件数が悪い場合には、これらの解法の収束状況はあまりよいとはいえない。特に、正規方程式を作りこれを解く場合には、元の行列の条件数を 2 倍にし、1 回の反復に必要な計算量も (1) 式に共役勾配法を適用したときより多い。さらに、丸め誤差の影響も受けやすいので、あまりおすすめできるアプローチとは言えない。しかし、共役勾配法自身はなかなか捨てがたいアルゴリズムである。そこで、なんとかよい修正方法を見つけ出せないだろうか。

そのアイデア一つは、共役勾配法を利用するのだが、行列の前処理にすこし工夫をこらしたものをを用いて共役勾配法の収束性を向上させることである。即ち、適切な行列の前処理を行って、係数行列 A のスペクトラムを改善し、それによって行列の条件数を減少させ、共役勾配法の収束をスピードアップすることである。

本報告では、従来から行列の前処理としてよく使われてきた不完全行列分解と共役勾配法の併用ではなく、次のような多項式に基づく行列の前処理と共役勾配法の併用を考えることにする (Adams[1], Dubois et al.[4], Johnson et al.[6], Saad[10] etc.).

$$\overline{Q}(A)Ax = \overline{Q}(A)b \quad (1.2)$$

ただし、 $\overline{Q}(A)$ は実係数の多項式で、これを前処理多項式と呼ぶことにする。このような行列による多項式の前処理で、(2) 式の右辺の係数行列 $Q(A)A$ が正定値行列になれば共役勾配法の収束が保証される。特に、行列 A がエルミート行列であり、 $Q(x)$ が実数を係数とする多項式であるならば、前処理した $Q(A)A$ もエルミート行列になる。

さらに、多項式の前処理を用いる利点の1つに、ベクトル計算機やパラレル計算機にその算法の実現のしやすさが上げられる (Adams[1], Johnson et al.[6], etc.).

2. 一般共役勾配法

行列の前処理は、次のように述べるができる。方程式 (1.1) を解く代わりに行列の前処理を行った次の方程式

$$BAx = Bb, \quad (2.1)$$

を解いても同値である。そこで、この方程式に共役勾配法を適用すると、次のアルゴリズムを得ることができる。これを一般共役勾配法 (Generalized Conjugate Gradient method) または、PCG 法という (Nodera[9]).

[ALGORITHM GCG]

- (1) 任意の初期値 x_1 を選び、残差ベクトル $r_1 = b - Ax_1$ および、方向ベクトル $p_1 = b_0 Br_1$ を計算する。ただし、行列 B は任意の対称正定値行列である。
- (2) 次の手順を繰り返す ($k = 1, 2, 3, \dots$).

$$x_{k+1} = x_k + a_k p_k, \quad (2.2a)$$

$$r_{k+1} = r_k - a_k A p_k, \quad (2.2b)$$

$$p_{k+1} = p_k + b_k B r_{k+1}, \quad (2.2c)$$

$$a_k = 1/(p_k, A p_k), \quad (2.2d)$$

$$b_k = 1/(r_{k+1}, B r_{k+1}). \quad (2.2e)$$

このアルゴリズムで重要な要因は、行列 B の選択方法であり、

- (1) 行列 BA は行列 A に対して固有値分布を改善し、 $\text{cond}(BA) \ll \text{cond}(A)$ を満足し、
- (2) $Br_{k+1} = d$ が簡単に計算できる、

ものでなければならない。ただし, $\text{cond}(C)$ は行列 C のスペクトル条件数を示すものとする。条件 (1) に関しては, 共役勾配法には次の性質が成立することによる。共役勾配法においては, 生成多項式としてチェビシェフ多項式を用いることによって, 近似解に対して次のような誤差の上限が得られることはよく知られている (Nodera[9])。

$$\|x_k - \hat{x}\|_C \leq 2 \left[\frac{\sqrt{(\text{cond}(C))} - 1}{\sqrt{(\text{cond}(C))} + 1} \right]^k \|x_1 - \hat{x}\|_C \quad (2.3)$$

ただし, $C = BA$ であり, \hat{x} は厳密解である。この式から共役勾配法の収束をつかさどるものの 1 つとして, 方程式の条件数が関与していることがあげられる。

行列 B の選択方法には, 従来, 次の (i) 行列分離, (ii) 行列分解, (iii) 不完全逆行列の 3 つに大別できる。(i) と (ii) に関しては現在までにいろいろな研究がすでに行われているのでその詳細についてはそれぞれの文献を参照してもらいたい (Meijrink et al.[6], Nodera[9], etc.)。本報告では, (iii) の多項式による不完全逆行列に基づいた行列の前処理について考えることにする (Dubois et al.[4], Johnson et al.[6], etc.)。

3. 多項式による前処理

多項式による行列の前処理の研究の歴史は以外と古く, 反復解法への応用は, 様々な文献に見いだすことができる (i.e. C. DeBoor et al.[3])。この前処理の第 1 の利点について述べると, そのアルゴリズムの単純さにあるのではないか。このアプローチの基本的な概念は, 不完全逆行列を共役勾配法の前処理として使うことであり, 1979 年に Dubois et al.[4]により提案された。

その基本的なアイデアは, 次のように述べることができる。行列 A の対角を全て 1 にスケールリングして, この行列を $A = I - F$ に分離する。ただし, F は対角がすべてゼロの対称行列である。通常, 行列 A は微分方程式の作用素を離散近似して構成されるものが多いので, F のスペクトル半

径は1よりも小さくなる。よって、行列 A の逆行列は、ノイマン級数展開を利用して、

$$B = (I - F)^{-1} = I + F + F^2 + F^3 + F^4 + \dots \quad (3.1)$$

書くことができる。不完全逆行列は、一言で述べれば、このノイマン級数を単純に途中で打ち切ったものを使うのである。この場合、行列 A が疎（スパース）であれば、行列 F も疎となり、不完全行列分解の場合と同様に多くの利点を持つ。しかし、この方法の弱点はなんといってもノイマン級数の収束が遅い点であり、充分使い物になる近似行列を得るのに、行列 F に関してたくさんの級数項をとらねばならないことにある。よって、不完全逆行列は、不完全行列分解よりも効果が得られないことが従来の数値実験からも報告されている。

不完全逆行列の収束性は、各係数にパラメータを導入することにより改善することが可能である。Johnson et al.[6]は、行列 A の近似逆行列として、次のようなパラメータをノイマン級数展開に導入した。

$$B_m = \sum_{j=0}^{m-1} k_j F^j. \quad (3.2)$$

これは行列 F に関連する多項式で行列 A を近似することである。この前処理行列 B によって、係数行列 A を前処理したものは

$$B_m A = \sum_{j=0}^{m-1} k_j A^{j+1} = \bar{Q}_m(A)A, \quad (3.3)$$

と書くことができる。また、 $B_m A$ は対称行列となり、そのスペクトラムは $\lambda_i p_m(\lambda_i)$, $i = 1, 2, \dots, n$ で与えられる。ただし、 λ_i は行列 A の固有値である。よって、 $B_m = \bar{Q}_m(A)$ であり、多項式である $Q_{m+1}(z) = z\bar{Q}_m(z)$ は $z = 0$ でゼロとなる。

言うまでもないが、ここで問題となるのは多項式の係数となるパラメータ $\{k_j\}$ の決定である。これは、共役勾配法が最適な収束をするよう

なものを使う必要がある。即ち、(3.2) 式の前処理によって、行列 $B_m A$ が正定値となり、さらにその固有値が 1 に近いものがよい。

このアイデアは、係数行列 A を $A = G - H$ に分離する場合にも適用することができる。この場合ノイマン級数展開に基づく前処理行列 B_m は、

$$B_m = \left(\sum_{j=0}^{m-1} k_j F^j \right) G^{-1}, \quad (3.4)$$

と書くことができる。ただし、 $F = G^{-1}H$ である。

このような多項式による前処理において、係数行列 A が対称でかつ正定値の場合に、Johnson et al.[6]は最適な多項式を minmax 近似によって決定するアルゴリズムを提案した。しかし、係数行列が対称ではあるが正定値でない場合には、彼等のアルゴリズムを適用することはできない。そこで次の節では、そのような係数行列に対して有効と思われる簡単な多項式の前処理について述べることにする。

4. 正定値でない対称行列の前処理

第 1 節で述べた多項式で前処理した (1.2) の方程式 $Q(A)Ax = Q(A)b$ を考える。ただし、係数行列 A は対称ではあるが正定値ではないものとする。前処理多項式は無限に構成することが可能だが、それを共役勾配法と併用した場合に計算量や収束の点から考えると、実際に有効となるものはそれほど多くない。誰もが考えつく基本的な多項式は、次の 3 つの型のものである。

$$\overline{Q}_a(z) = (k_a + z) \quad (4.1a)$$

$$\overline{Q}_b(z) = z(k_b - z) \quad (4.1b)$$

$$\overline{Q}_c(z) = z(k_c^2 - z^2) \quad (4.1c)$$

これらの多項式によって前処理した係数行列は、 $\overline{Q}(A)A$ となりこれを $Q(A)$ とおくことにする。いうまでもないが、多項式 $Q(z)$ は $Q(0) = 0$

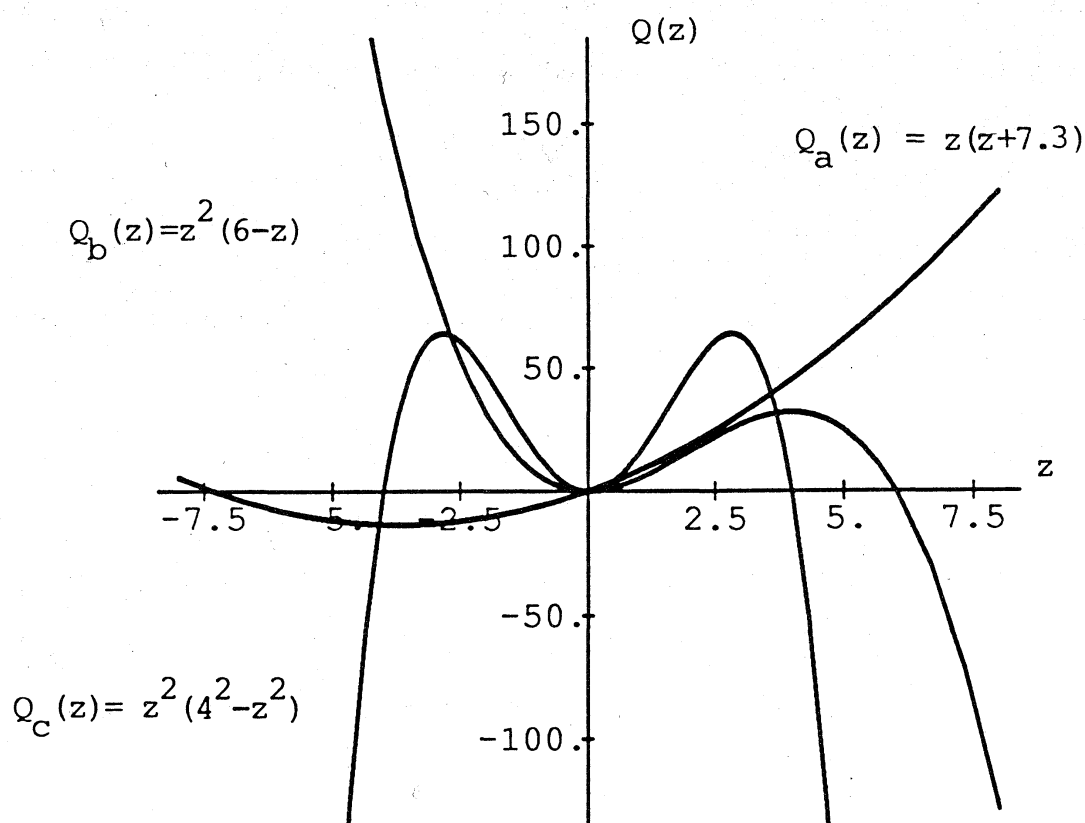


図 4.1 前処理多項式の例

となり、この式は原点を通過し、その次数は 2 に等しいかそれ以上である ($Q(z) \geq 2$)。これは、この前処理によって正定値性を作り出すことに起因する。また、当然のことながら、多項式の係数 $\{k_j\}$ は、方程式の条件数が最小となるように決定する必要がある。即ち、 $\text{cond}(Q(A))$ が $\text{cond}(A^2)$ よりも充分小さくなるように決めなければならない。これは、共役勾配法を使って正規方程式を解く場合よりも、この前処理が有効に働く必要があることを示している。

次にこの前処理の有効性を示すために、新しい条件数を定義する。

定義. $R = [a, b] \cup [c, d]$ を実軸上の部分集合とする。この部分集合 R には行列 A のスペクトラム $\sigma(A)$ を含むものとする。さらに、多項式 $Q(z)$ は

$\forall z \in R$ に対して, $Q(z) > 0$ を満足するものとする. ここで, 行列 $Q(A)$ の事実上のスペクトラル条件数 (疑似的なスペクトラル条件数) を

$$\overline{\text{cond}}(Q(A)) = \frac{\max_{z \in R} Q(z)}{\min_{z \in R} Q(z)} \quad (4.2)$$

と定義することにする. ここで, もし $a, b, c, d \in \sigma(A)$ ならば, 区間 R 上で正となる行列 $Q(A)$ に対して, 事実上のスペクトラル条件数 $\overline{\text{cond}}(Q(A))$ はただ 1 つ決定できる.

この定義により, 次のような定理が得られることになる.

定理 4.1. 行列 A は対称ではあるが, 正定値ではないものとする. また, そのスペクトラム $\sigma(A)$ は, $\sigma(A) \subset R = [a, b] \cup [c, d], b < 0 < c$ を満足するものとする. また, 多項式 $Q(z)$ は, $\forall z \in R$ において $Q(z) > 0$ となるものとする. (真の) スペクトラル条件数と事実上のスペクトラル条件数の間には次の関係が成立する.

$$\text{cond}(Q(A)) \leq \overline{\text{cond}}(Q(A)) \quad (4.3)$$

証明: 仮定から, 行列 A のスペクトラムに関して, $\sigma(A) \subset R$ が成立し多項式 $Q(z)$ も R 上で正なので ($\forall z \in R, Q(z) > 0$),

$$\min_{z \in \sigma(A)} Q(z) \leq \min_{z \in R} Q(z) \text{ および } \max_{z \in \sigma(A)} Q(z) \leq \max_{z \in R} Q(z)$$

が成立する. よって,

$$\begin{aligned} \text{cond}(Q(A)) &= \frac{\max_{z \in \sigma(A)} Q(z)}{\min_{z \in \sigma(A)} Q(z)}, \\ &\leq \frac{\max_{z \in R} Q(z)}{\min_{z \in R} Q(z)}, \\ &= \overline{\text{cond}}(Q(A)). \blacksquare \end{aligned}$$

この定理から言えることは, 前処理多項式の次数の決定は, 行列 A のスペクトラムを含む R の区間に依存することになる. 例えば, (4.1a) 式について, 次の系 4.2 が成立する.

系 4.2. 行列 A は対称ではあるが, 正定値でないものとする. また, そのスペクトラム $\sigma(A)$ は次の条件を満たすものとする.

$$\sigma(A) \subset R = [a, b] \cup [c, d]$$

ただし, $-a \leq d, b < 0 < c$ である. 前処理多項式 $Q(z)$ として R 上で正となる $Q_k(z) = z(z - k)$ を考える. ただし, $k \leq a + d$ である. ここで, もし $k_0 = b + c \leq a + d$ ならば

$$\overline{\text{cond}}(Q_{k_0}(A)) \leq \overline{\text{cond}}(Q_k(A)) \quad (4.4)$$

が成立する. さらに, $a, b, c, d \in \sigma(A)$ ならば

$$\text{cond}(Q_{k_0}(A)) \leq \text{cond}(Q_k(A)) \quad (4.5)$$

が成立する.

証明: 省略 (Nandakumar[8]を参照せよ).

次のような簡単な例をここで考えてみよう.

[例 1] 行列 A のスペクトラムが 2 つの区間 $[-9.0, -6.0] \cup [0.2, 3.2]$ に含まれているものとする. ここで, この行列 A のスペクトラル条件数と事実上のスペクトラル条件数を計算すると次のようになる.

$\text{cond}(A)$	45
$\text{cond}(A^2)$	2025
$\overline{\text{cond}}((5.8I + A)A)$	24

この例からも理解できるように, 対称行列ではあるが正定値ではない係数行列に対してもそのスペクトラムが含まれるおおまかな区間の情報が少しでも入手できるならば, 多項式の前処理の有効性は高いように思われる. しかし, 実際には, 行列 A のスペクトラムはわからないので, この

ような多項式の推定は不可能に近い。また、他の前処理と同様に不確定な要因もたくさん残されている。

超高速の演算スピードを有するスーパーコンピュータや並列計算機において PCG 法を実行する場合に、演算速度をダウンさせるため従来ネットワークとされていた前処理行列の前身および後退代入の操作が、多項式による前処理を用いれば必要でなくなり (行列) \times (ベクトル) の計算だけで済んでしまう利点がある。即ち、反復法でいえばヤコビ法のような単純計算をすることになる。しかし、前にも述べたが、アルゴリズムの収束に関して少し問題が残らないわけではないので、これからの研究に委ねる所が多いように思われる。

最後に、多項式による前処理とそうでない不完全行列分解とを使った前処理の融合を考えてみよう。このような前処理の存在は、例えば、Adams[1]がすでに示唆しており、その前処理行列は $B = Q(MA)M$ で与えることができる。ただし、 $Q(z)$ は多項式であり、 M は行列 A の不完全行列分解による近似逆行列を使うことになる。このような多項式の前処理と多項式によらない前処理を融合することによって共役勾配法の収束を加速することができる。

REFERENCES

1. L. Adams, *m-Step preconditioned conjugate gradient methods*, SIAM J. Sci. Stat. Comput. **6** (1985), 452-463.
2. R. Chandra, S.C. Eisenstat and M.R. Scultz, *The modified conjugate residual method for partial differential equations*, Advances in Computer Methods for Partial Differential Equations II (1977), 13-19.
3. C. De Boor and J.R. Rice, *External polynomials with applications to Richardson iteration for indefinite systems*, Math Res. Center, Univ. of Wisconsin, Madison (1980).
4. P.F. Dubois, A. Greenbaum and G.H. Rodrigue, *Approximating the inverse of a matrix for use in iterative algorithms on vector processors*, Computing **22** (1979), 257-268.
5. M. R. Hestenes and E. Stiefel, *Methods of conjugate gradients for solving linear systems*, J. Res. Nat. Bur. Stand. (1952), 409-436.
6. O.G. Johnson, C.A. Michell and G. Paul, *Polynomial preconditioners for conjugate gradient calculations*, SIAM Numer. Anal. **20** (1983), 362-376.

7. J.A. Meijerink and H.A. van der Vorst, *An iterative solution method for linear systems of which the coefficient matrix is a symmetric M-matrix*, Math. Comp. **31** (1977), 148-162.
8. N.R. Nandakumar, *Polynomial preconditioning of symmetric indefinite systems*, Univerity of Illinois, CSRD No.580 (1986).
9. T. Nodera, *PCG method for large sparse linear systems (in Japanese)*, Seminar on Mathematical Science **7** (1983).
10. Y. Saad, *Iterative solution of indefinite symmetric linear systems by the methods using orthogonal polynomials over two disjoint intervals*, SIAM **20** (1983), 784-811.